

Monty Python Reloaded

YOU WON'T

BELIEVE WHAT

THIS GROUP DID 

Monty Python Reloaded

YOU WON'T
BELIEVE WHAT
THIS GROUP
Clickbait detected



MONTy PythoN REloaded

Clickbait Classifier

David KUPERT, Felix TEUTSCH, JONAS BEER

Agenda

1. Topic
2. The motivation
3. State of the Art
4. Methodology
5. Experimental Results
6. Explainability
7. Contribution

Topic - Clickbait

Clickbait refers to headlines that entice readers to click on a link, often leading to content of bad quality or relevance

Topic - Clickbait detection

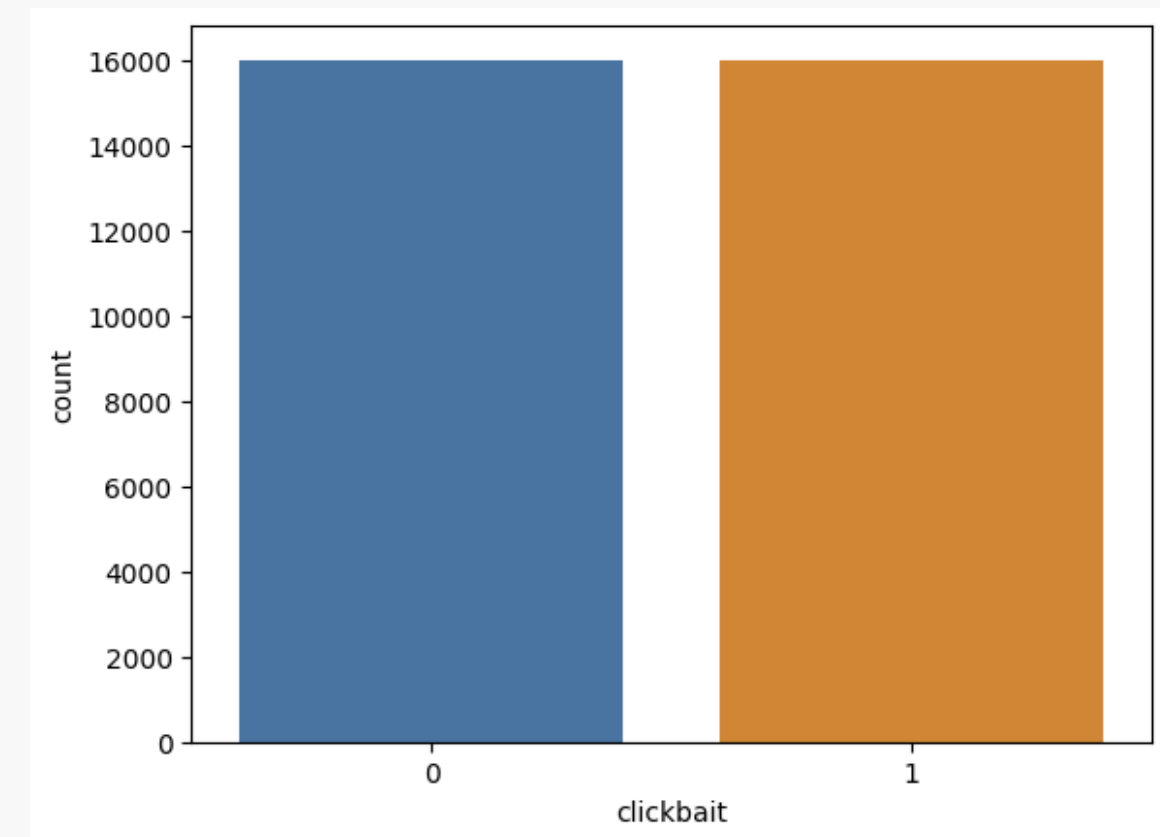
Clickbait detection is important to ensure:

- information integrity
- user experience
- digital literacy

Topic - Dataset

Our dataset consists of:

- clickbait headlines (Buzzfeed, ViralNova, ...)
- non-clickbait headlines (WikiNews, NY Times, The Guardian)
- 0 or 1 depending on the category



The Goal / Motivation

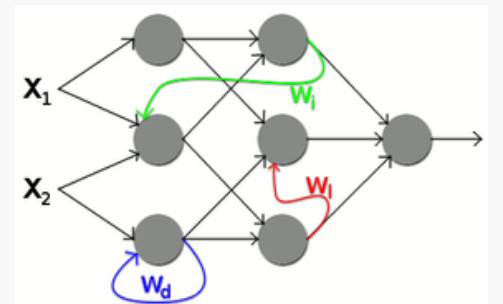


- Understanding the Methods behind Clickbait detection
- Personal issues with clickbait
- Improve Classifier



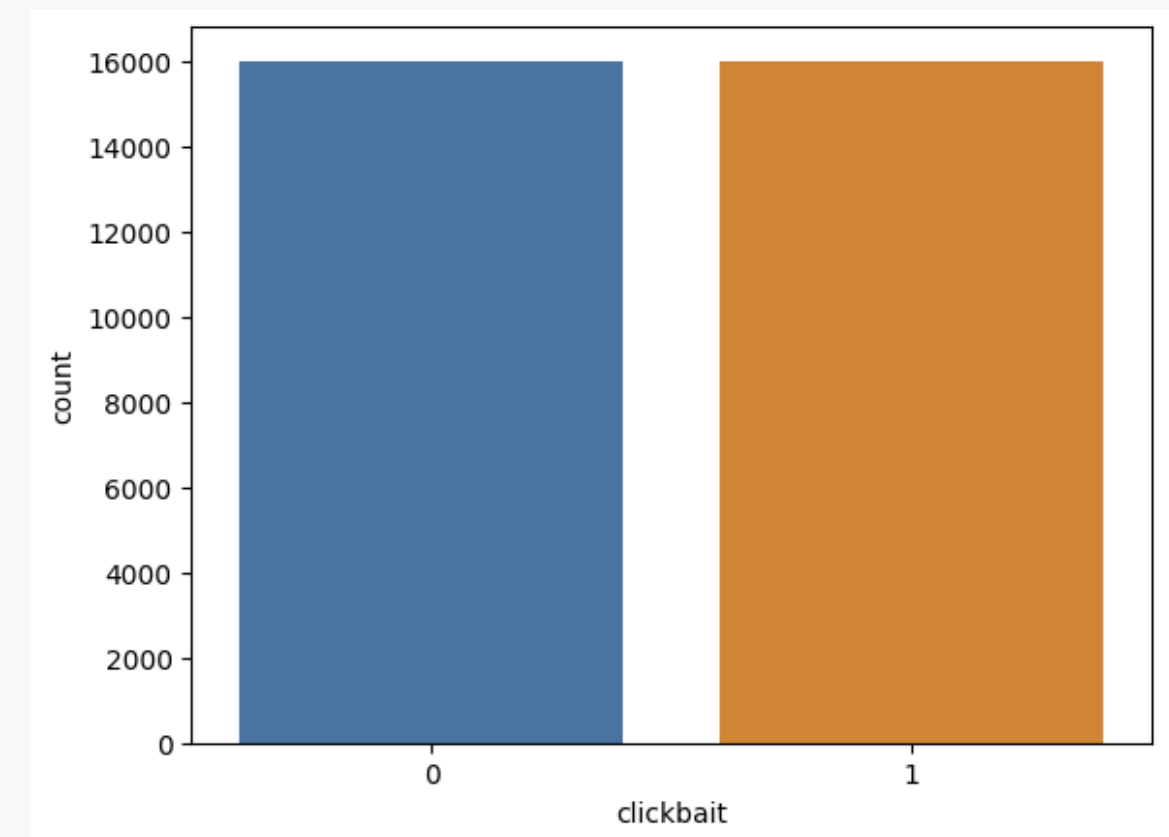
State of the Art

- Clickbait Detection via Large Language Models
 - “Experimental results show that LLMs cannot achieve the best results compared to the state-of-the-art deep and fine-tuning PLMs methods”, used ChatGPT 3.5
- Clickbait Detection with Style-aware Title Modeling and Co-attention
 - Used style-aware title modelling and co-attention to improve clickbait detection
- We used Neural Networks to Detect Clickbaits: You won't believe what happened Next!
 - Created a high-performing classifier (F1=98.19%), used RNN
- Stop Clickbait: Detecting and Preventing Clickbaits in Online News Media
 - Used RandomForest, SVM, Decision Tree
 - SVM was the best Classifier with F1=93%



Methodology

- Exploratory data analysis:
 - data consist of 32k headlines
 - 50% clickbait, 50% non-clickbait → balanced dataset
 - no missing values



Methodology

- Data Preparation
 - removing stopwords / special characters
 - removing punctuation
 - converting all headlines to lowercase



```
#Demonstrating the Text Preprocessing
```

```
print(CleanText('This is a headline that is totally not clickbait, unbelievable!!!'))
```



```
headline totally clickbait unbelievable
```

Methodology

- Tried 5 different classifiers we learned in our course
 - Logistic Regression
 - SGD Classifier (Stochastic Gradient Descent)
 - KNN
 - Decision Tree
 - Random Forest

EXPERIMENTal RESULTS

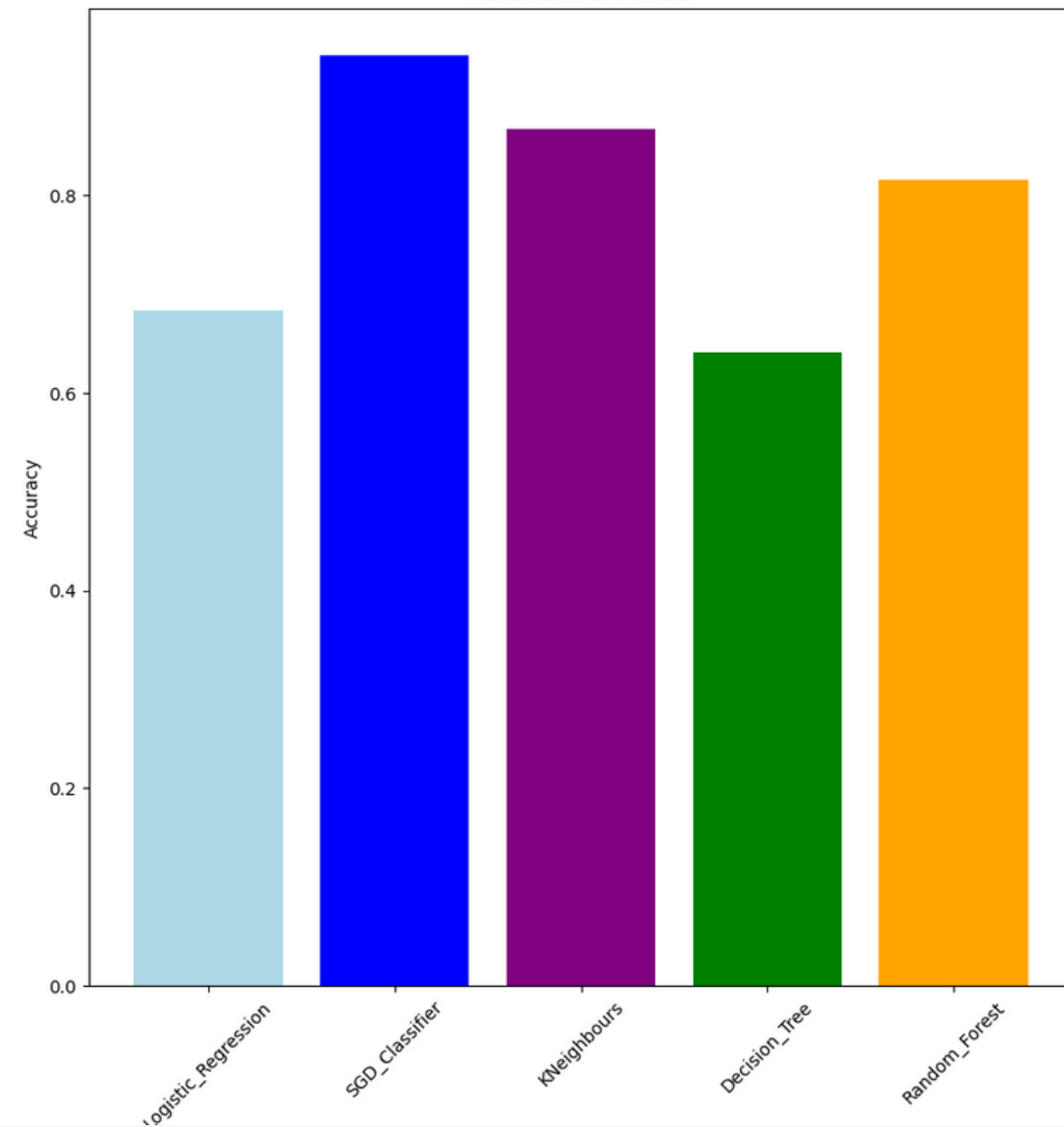
- Chosen Classifier
 - best performing overall classifier: SGD
 - accuracy: 94.23%
 - precision: 94.23%
 - recall: 94.23%
 - F1: 94.23%
 - ROC: 94%

```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  PORTS

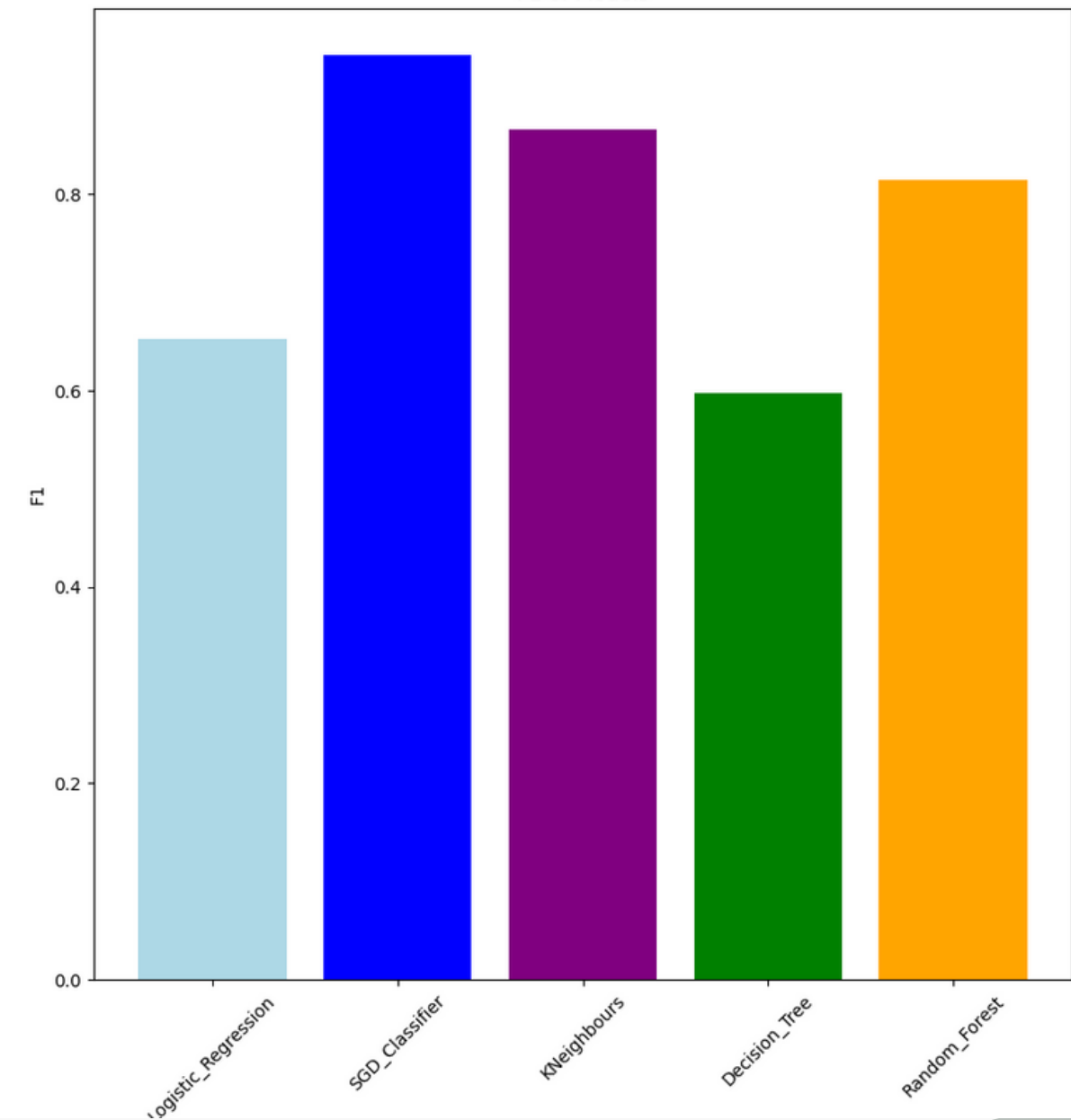
-----
The accuracy score of Logistic_Regression is 0.68390625
The precision score of Logistic_Regression is 0.8038749216979292
The recall score of Logistic_Regression is 0.68390625
The F1 score of Logistic_Regression is 0.6526342292050222
-----
The accuracy score of SGD_Classifier is 0.94234375
The precision score of SGD_Classifier is 0.9423453726331448
The recall score of SGD_Classifier is 0.94234375
The F1 score of SGD_Classifier is 0.9423443541720207
-----
The accuracy score of KNeighbours is 0.8678125
The precision score of KNeighbours is 0.8884510141362861
The recall score of KNeighbours is 0.8678125
The F1 score of KNeighbours is 0.8656224527775089
-----
The accuracy score of Decision_Tree is 0.6415625
The precision score of Decision_Tree is 0.77234808479192
The recall score of Decision_Tree is 0.6415625
The F1 score of Decision_Tree is 0.5972921556322573
-----
The accuracy score of Random_Forest is 0.8153125
The precision score of Random_Forest is 0.8263332122297018
The recall score of Random_Forest is 0.8153125
The F1 score of Random_Forest is 0.8142179676354891
```

EXPERIMENTAL RESULTS

Accuracies of Models

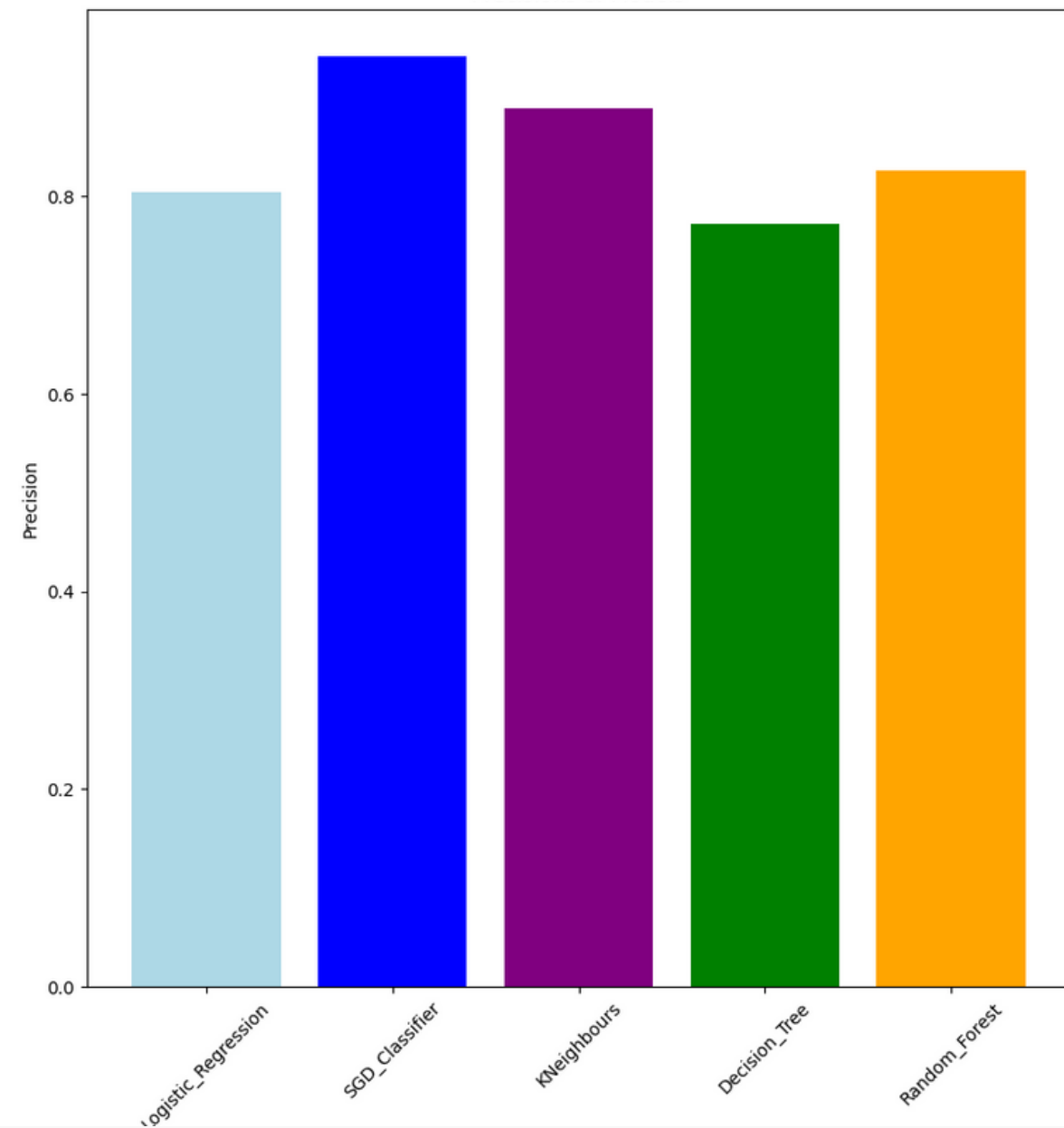


F1 of Models

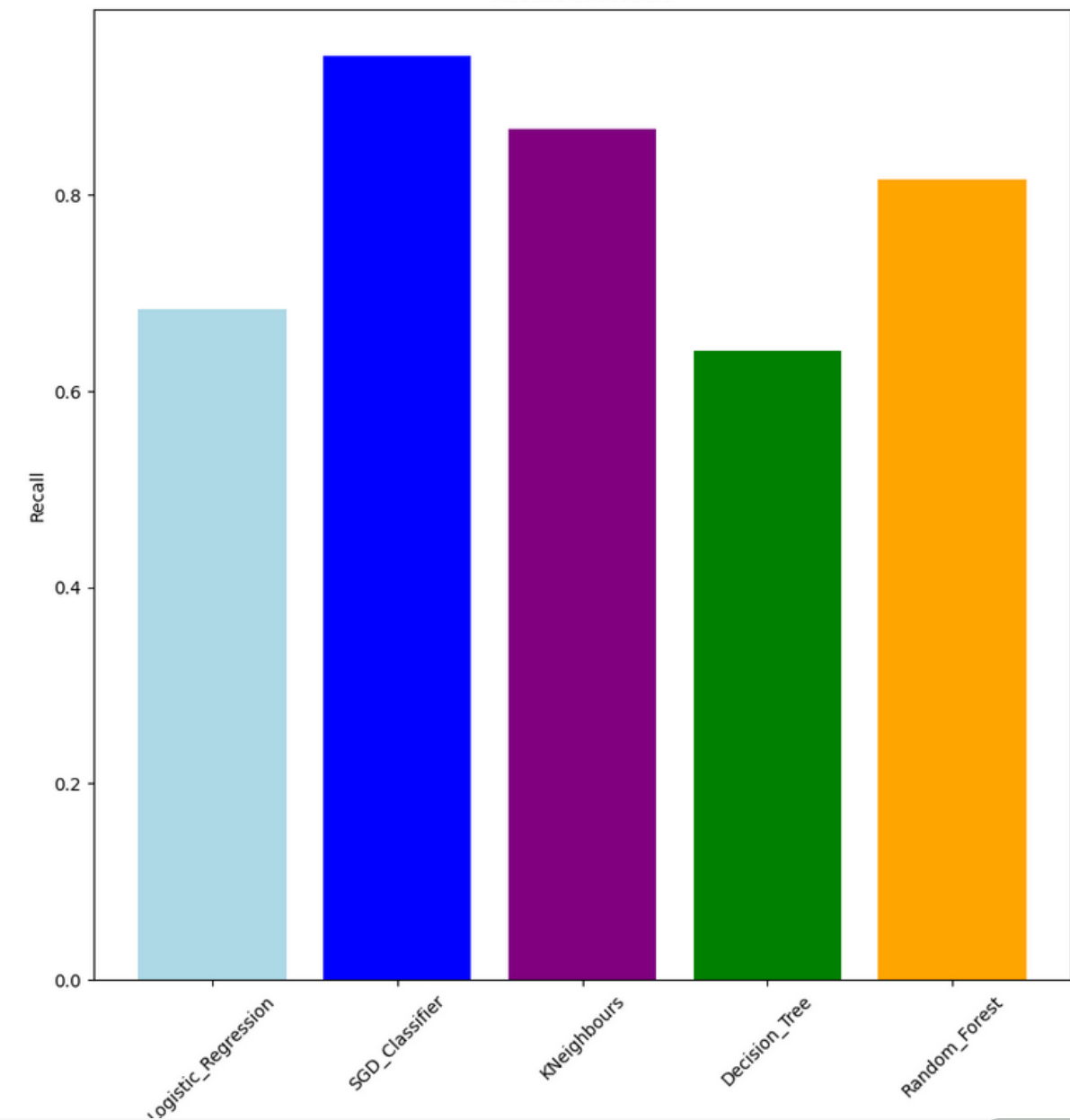


EXPERIMENTal RESULTS

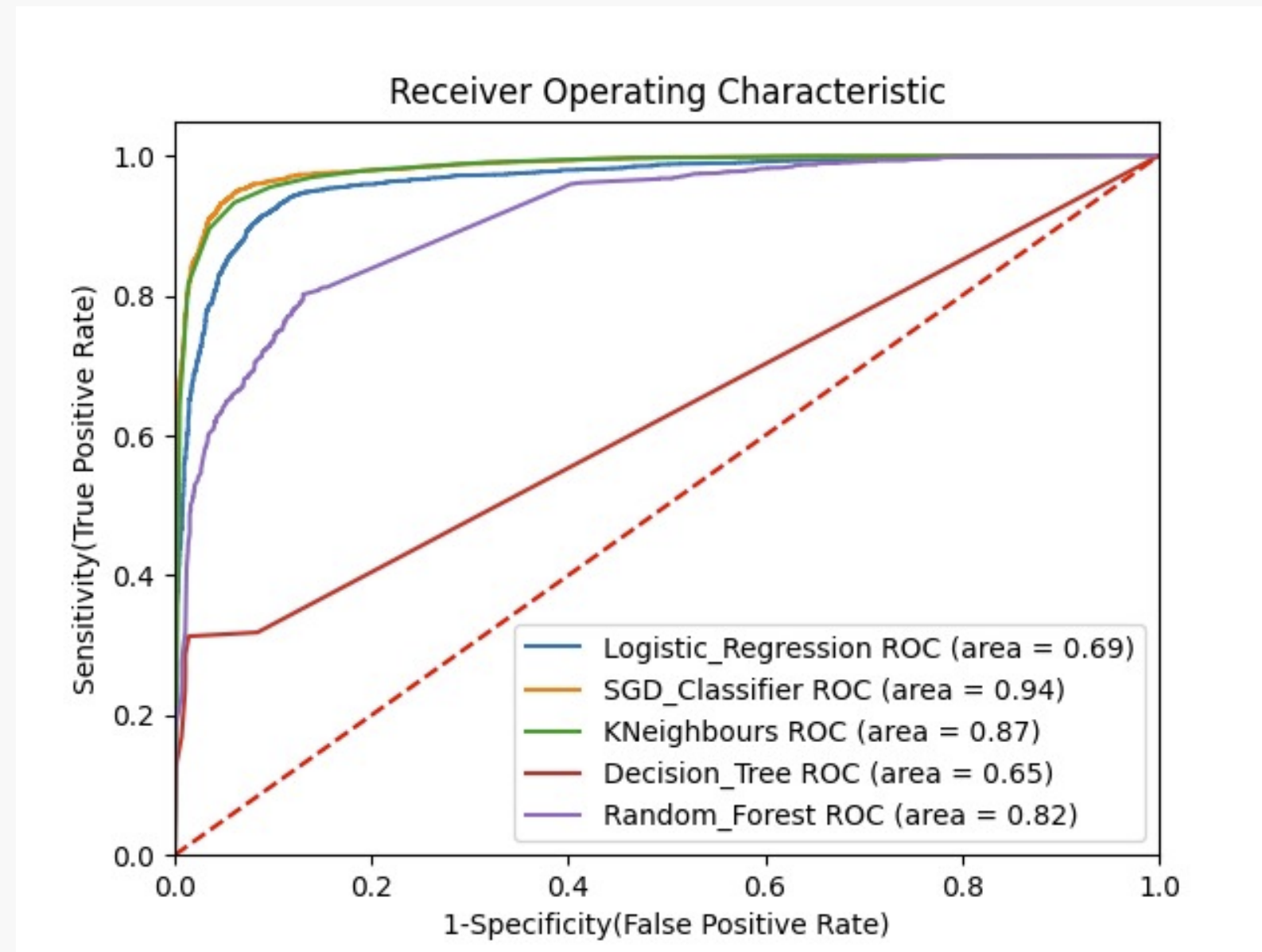
Precisions of Models



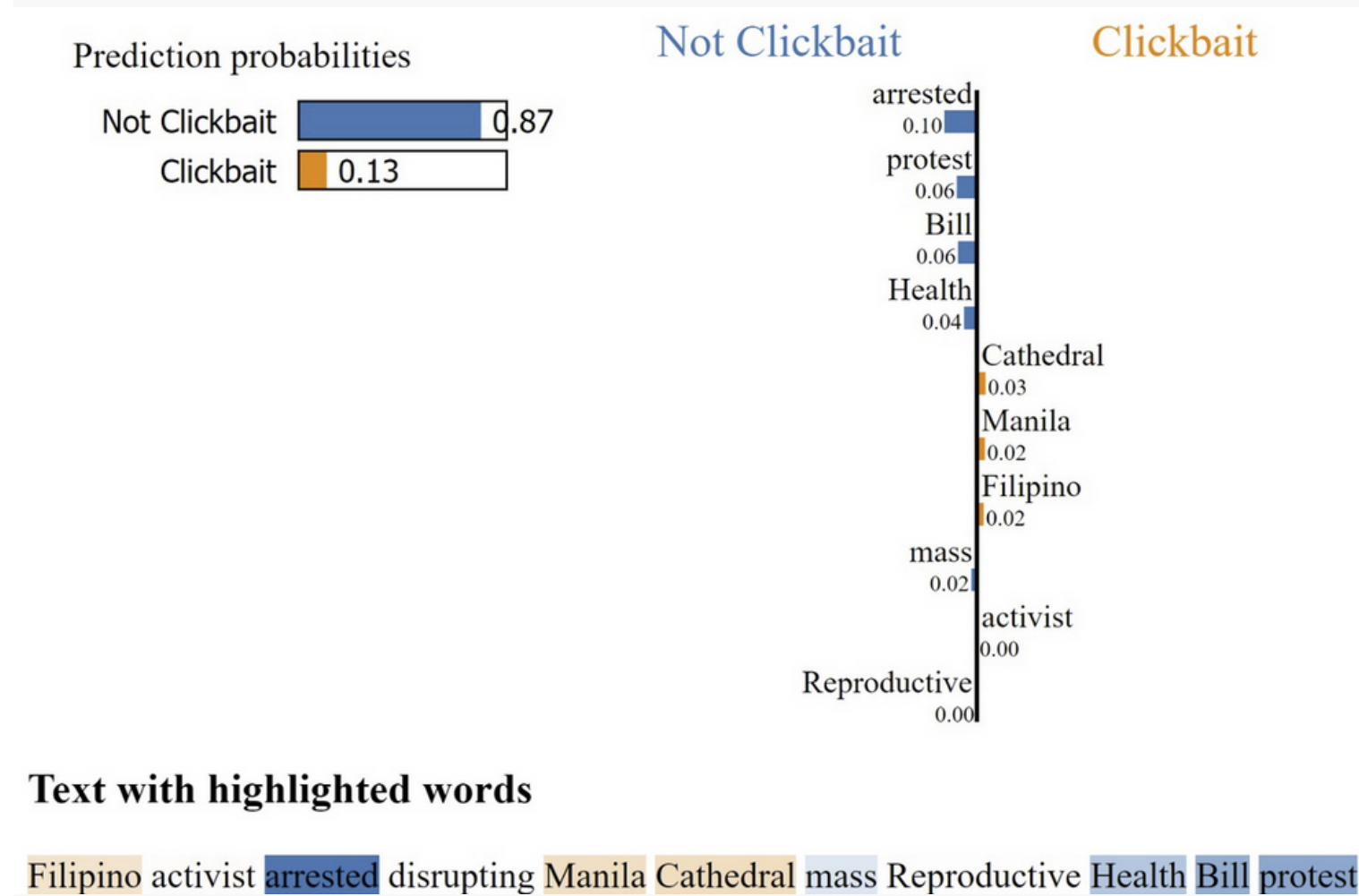
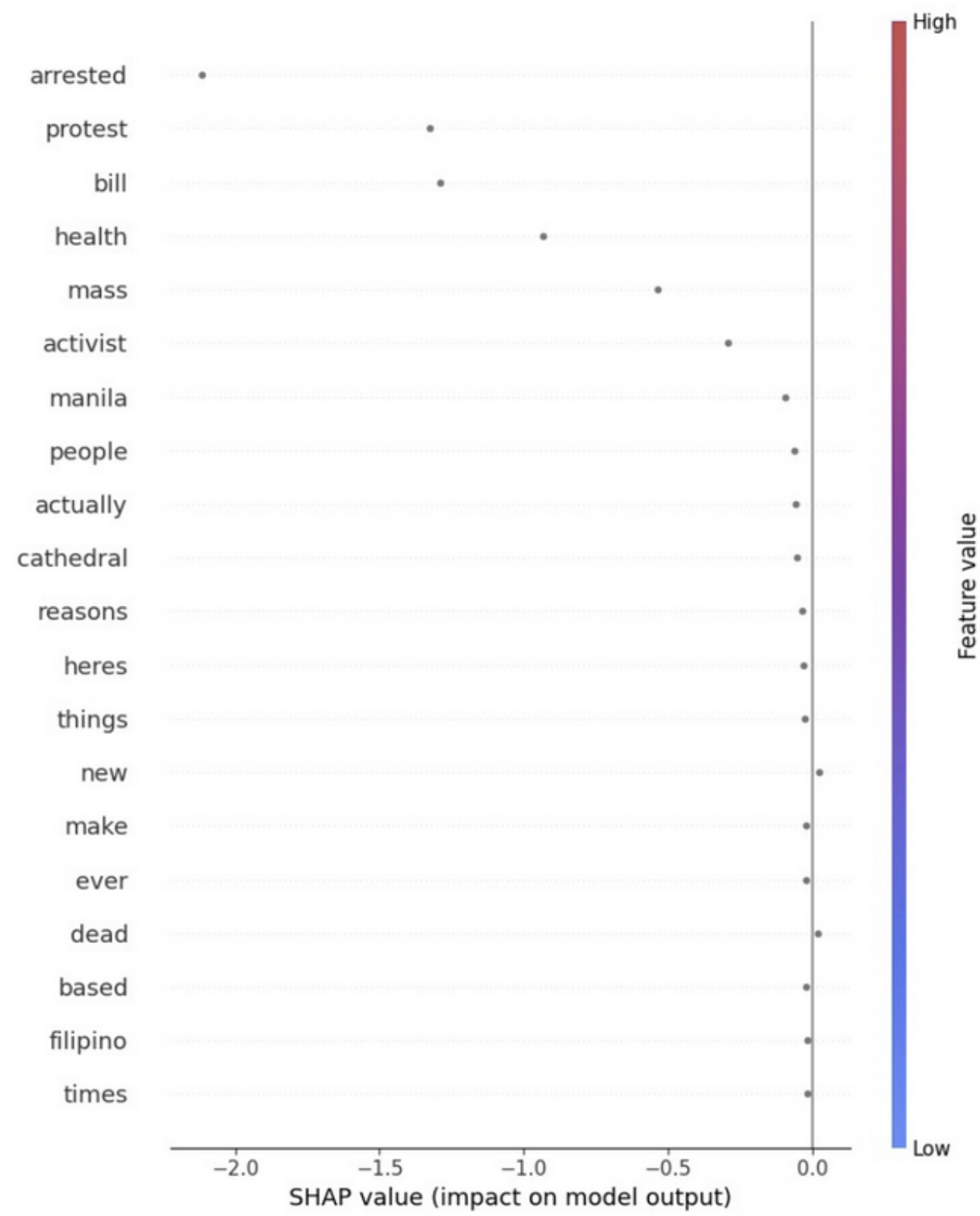
Recalls of Models



EXPERIMENTal RESULTS

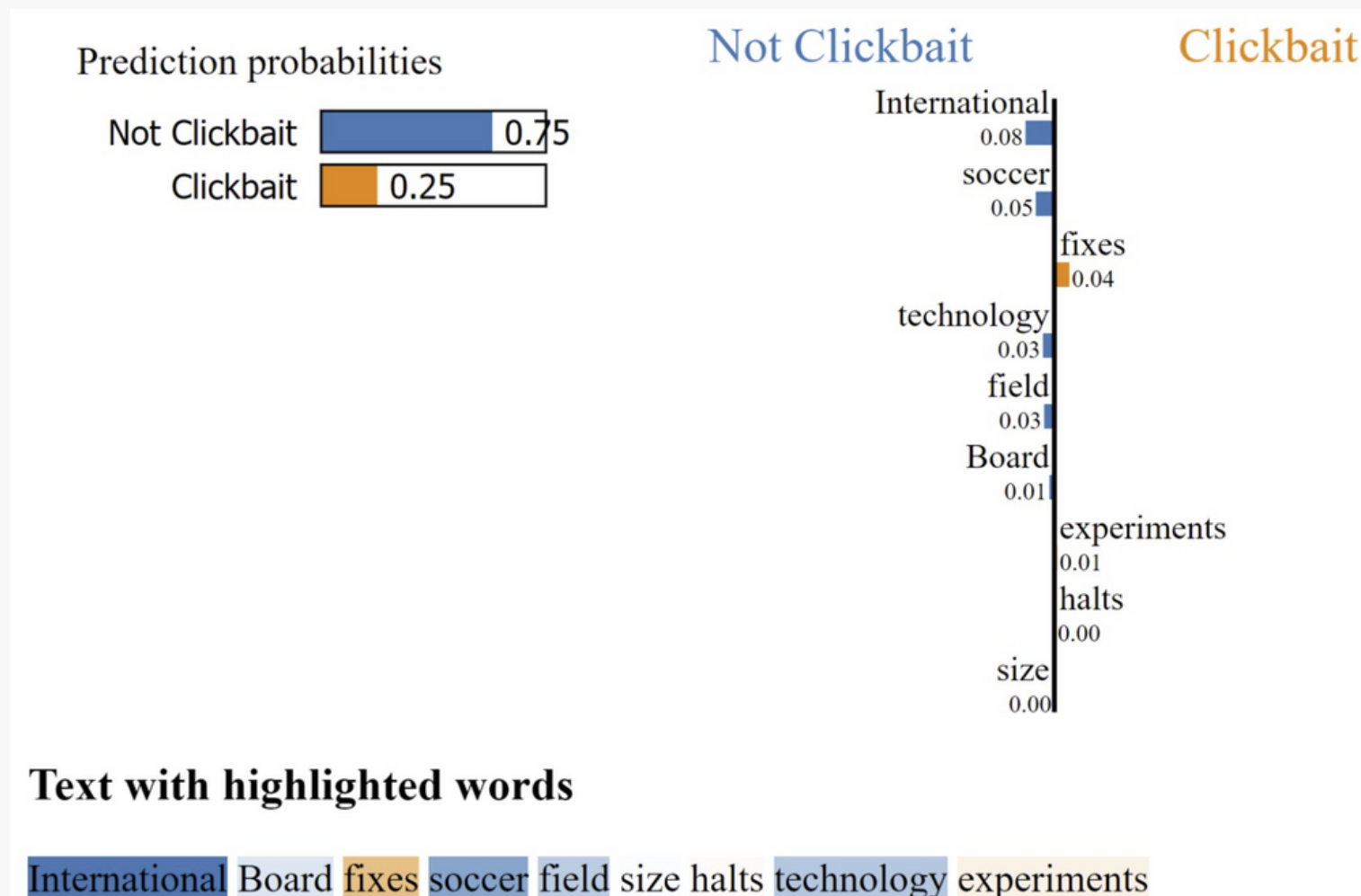
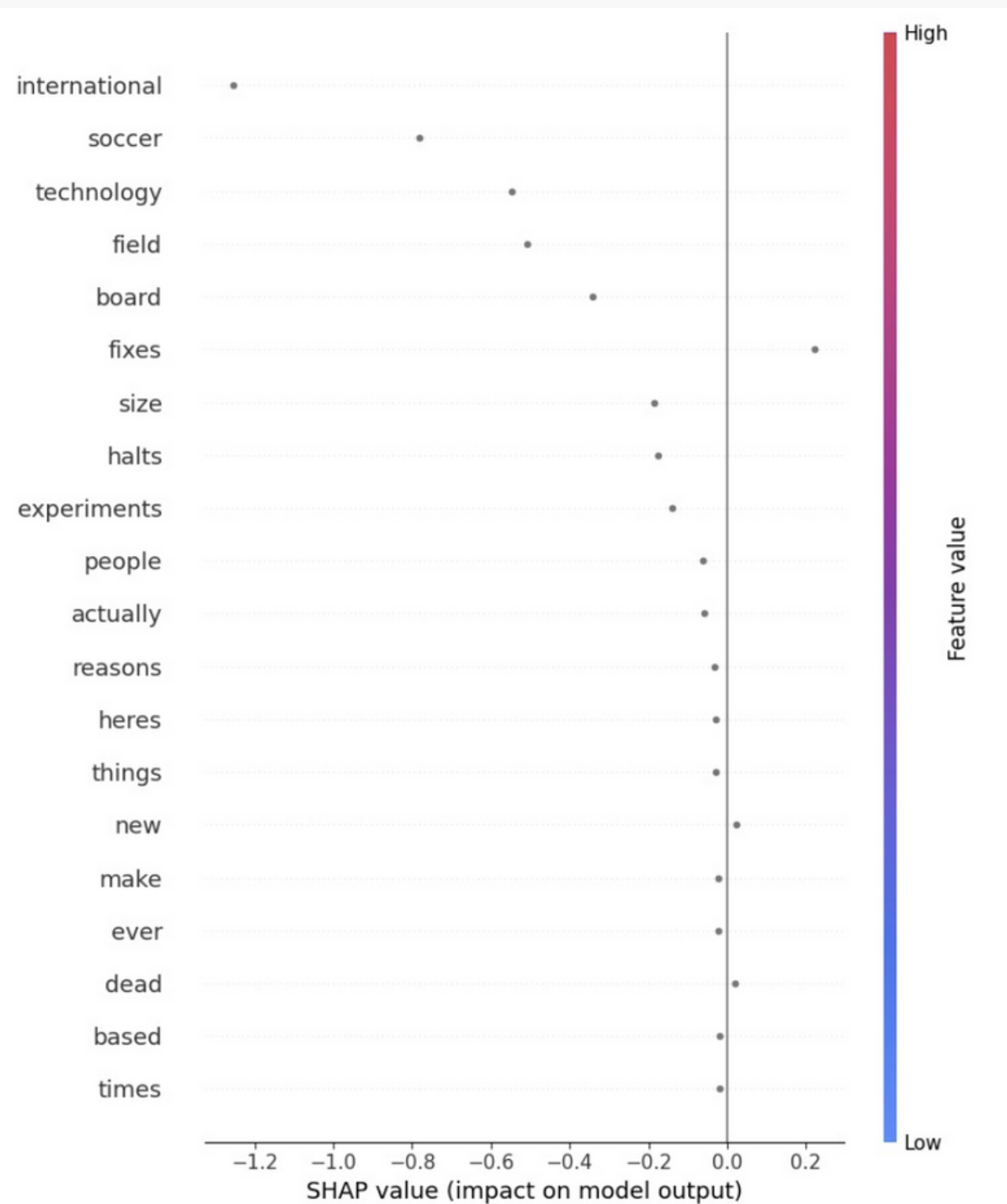


Explainability - Felix



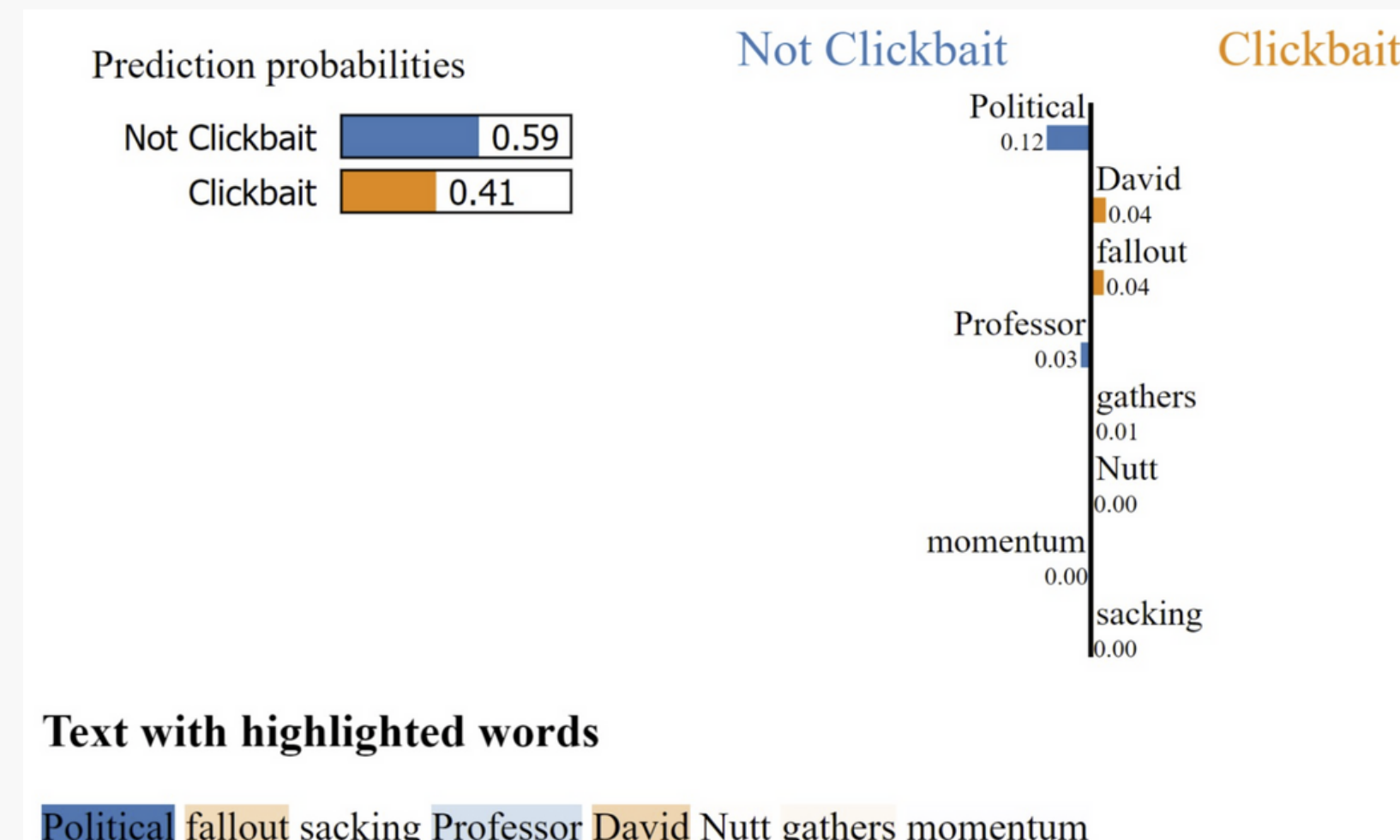
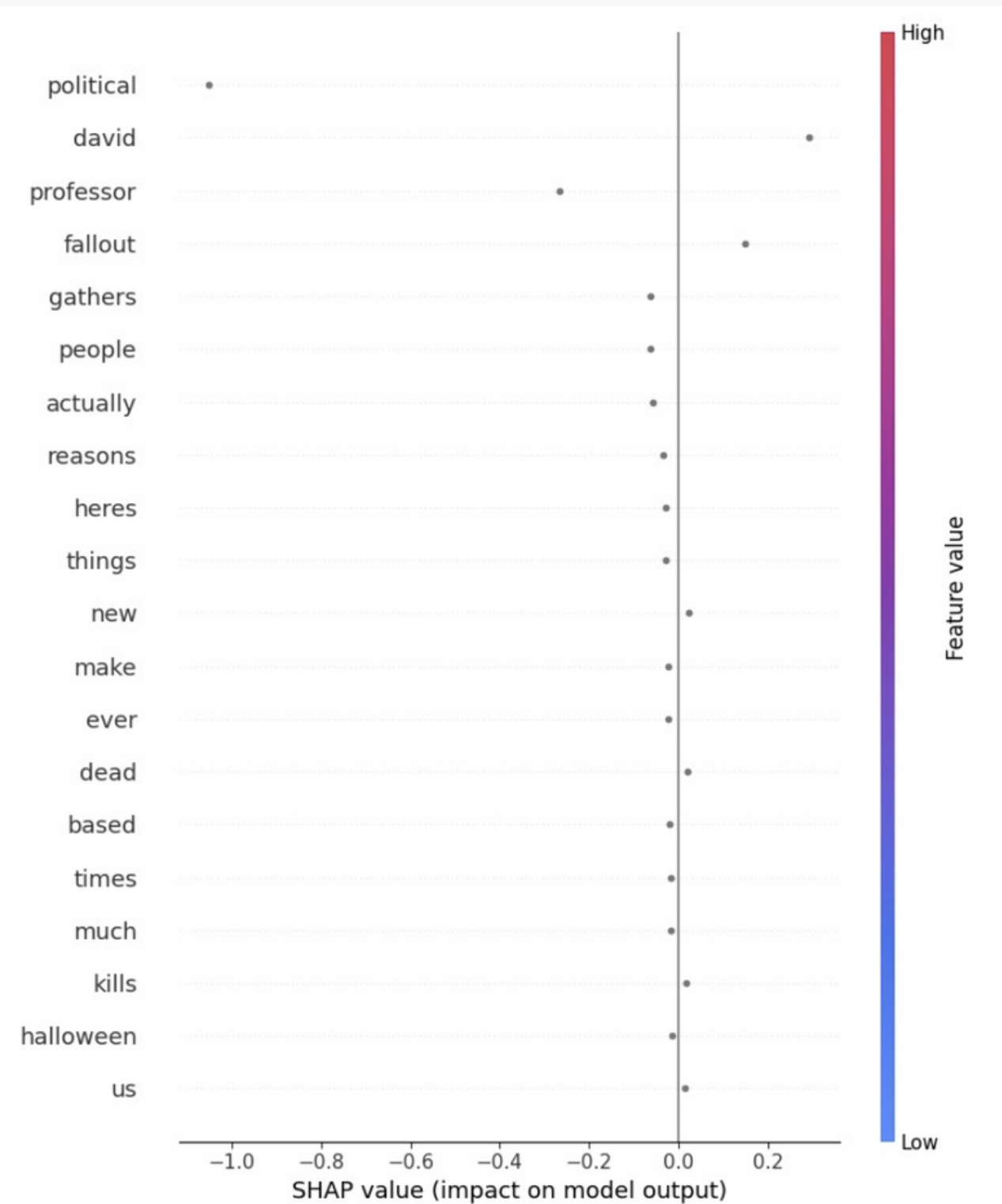
Filipino activist arrested
disrupting Manila Cathedral
mass Reproductive Health Bill protest

Explainability - JONAS



International Board fixes soccer field size
halts technology experiments

Explainability - David



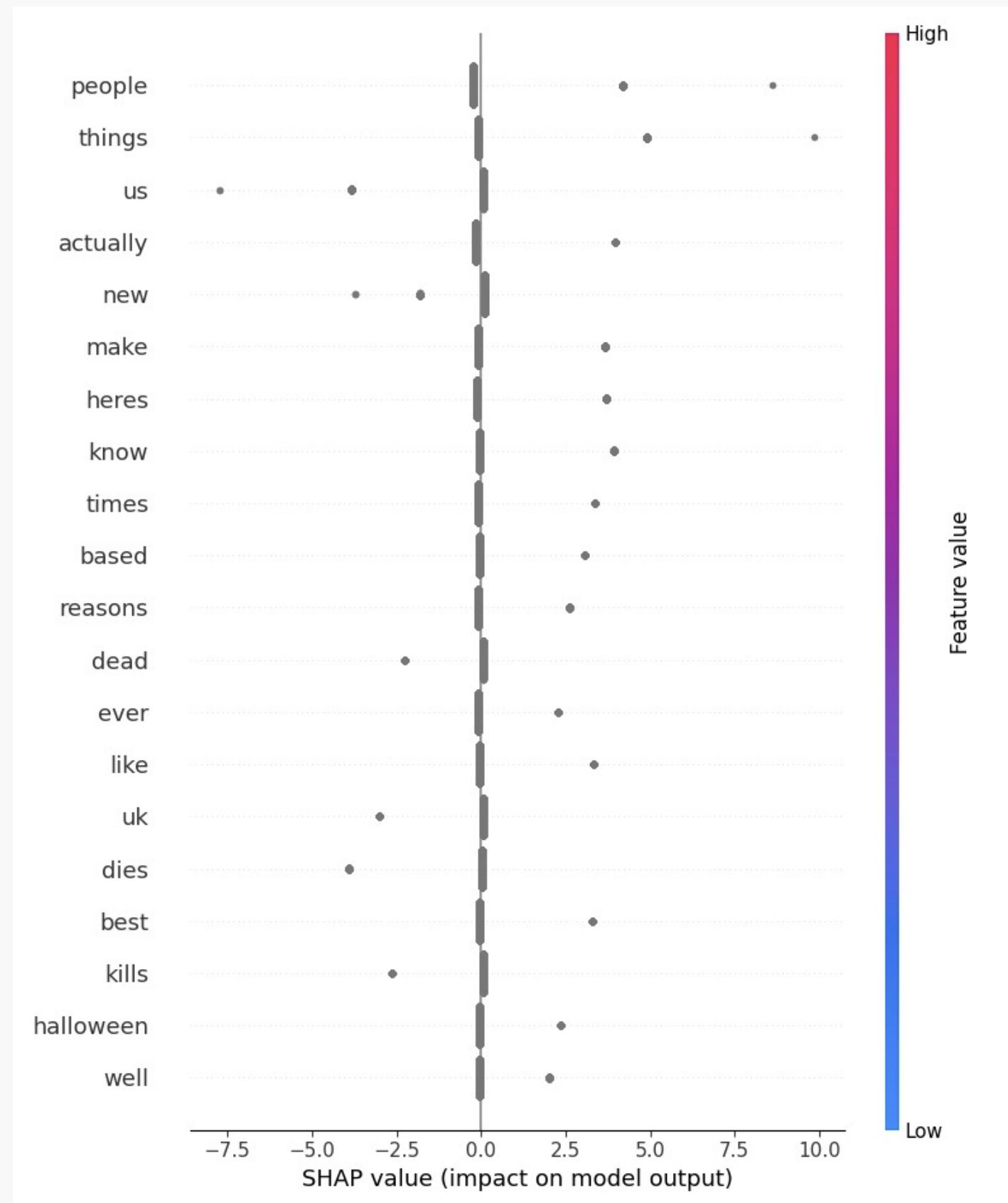
Political fallout sacking Professor David Nutt
gathers momentum

Explainability

Our group interpretation:

- Words that are often used in a sensationalized context are **more** likely to be clickbait
- action-oriented words are **more** likely to be clickbait (call to actions)
- “boring” and serious words are **less** likely to be clickbait
- we prefer **LIME** over **SHAP**

Explainability of most used words



CONTRIBUTION

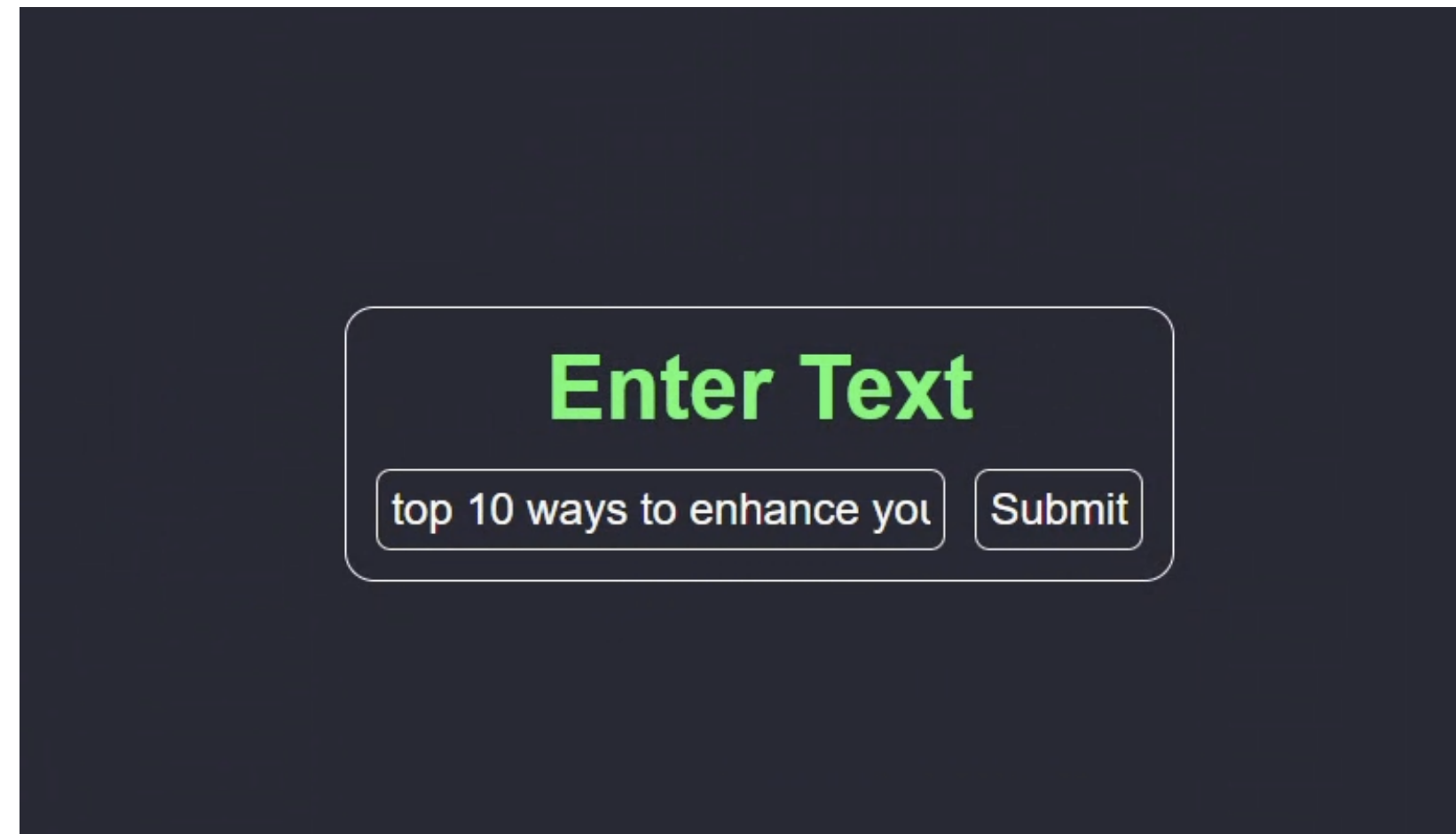
TOP 3 WAYS TO **IMPROVE**
PERFORMANCE BY 0.26% (you won't
believe number 3!!!)

- Adding numbers - David
- Not converting headlines to lowercase - Jonas
- Adding exclamation and question marks - Felix



CONTRIBUTION

Web demo of a live Clickbait classifier

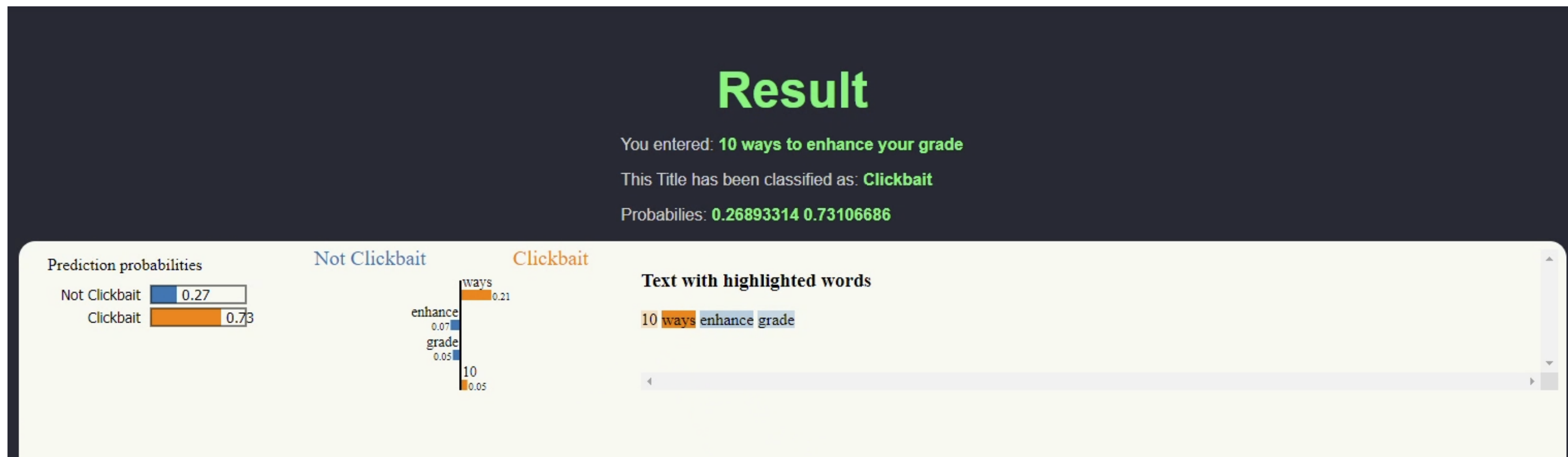
A screenshot of a web interface with a dark blue background. In the center, there is a rounded rectangle containing the text "Enter Text" in green. Below this, there is a text input field containing the text "top 10 ways to enhance you" and a "Submit" button to its right.

Enter Text

top 10 ways to enhance you

CONTRIBUTION

Web demo of a live Clickbait classifier



QUESTIONS?

Q & (maybe) A

All

Thank you for
your Attention